The Electronic Library

Emerald Article: Developing a Taiwan library history digital library with reader knowledge archiving and sharing mechanisms based on the DSpace platform

Chih-Ming Chen, Yong-Ting Chen, Chin-Ming Hong, Chin-Wen Liao, Chia-Meng Huang

Article information:

To cite this document:
Chih-Ming Chen, Yong-Ting Chen, Chin-Ming Hong, Chin-Wen Liao, Chia-Meng Huang, (2012), "Developing a Taiwan library history digital library with reader knowledge archiving and sharing mechanisms based on the DSpace platform", The Electronic Library, Vol. 30 Iss: 3 pp. 426 - 442

Permanent link to this document:
http://dx.doi.org/10.1108/02640471211241681

Downloaded on: 06-06-2012

References: This document contains references to 10 other documents

To copy this document: permissions@emeraldinsight.com

Access to this document was granted through an Emerald subscription provided by Emerald Author Access

For Authors:
If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service. Information about how to choose which publication to write for and submission guidelines are available for all. Additional help for authors is available for Emerald subscribers. Please visit www.emeraldinsight.com/authors for more information.

About Emerald  www.emeraldinsight.com
With over forty years’ experience, Emerald Group Publishing is a leading independent publisher of global research with impact in business, society, public policy and education. In total, Emerald publishes over 275 journals and more than 130 book series, as well as an extensive range of online products and services. Emerald is both COUNTER 3 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.
Developing a Taiwan library history digital library with reader knowledge archiving and sharing mechanisms based on the DSpace platform

Chih-Ming Chen and Yong-Ting Chen
Graduate Institute of Library, Information and Archival Studies, National Chengchi University, Wenshan District, Taipei City, Taiwan

Chin-Ming Hong
Department of Applied Electronics Technology, National Taiwan Normal University, Taipei, Taiwan

Chin-Wen Liao
Department of Industrial Education and Technology, National Changhua University of Education, Changhua City, Taiwan, and

Chia-Meng Huang
ADLINK Technology Inc., Taipei, Taiwan

Abstract

Purpose – This work seeks to present a reading annotation and knowledge sharing tool, which can annotate a web page with HTML format archived by the Taiwan libraries’ history digital library based on Web 2.0 technologies.

Design/methodology/approach – This work adopted DSpace, an open-source institutional repository system, to implement a Taiwan Digital Library History Library with the reading annotation tool for knowledge archiving and sharing services. A quasi-experimental design method was employed to randomly assign participants to an experimental group and control group to evaluate differences in the reading performance of learners who used the proposed annotation system. A statistical analysis scheme was employed to evaluate differences in learning performance of learners while reading and learning with the proposed annotation tool.

Findings – The paper finds that annotated digital material provides useful knowledge to readers. The values to those annotating and subsequent readers are the acquisition of in-depth knowledge and efficient reading. Additionally, the effect on digital libraries is that digital library content grows dynamically as readers contribute knowledge. More importantly, annotated information from different readers has very high potential for the discovery of value-added knowledge utilizing data mining techniques.

Originality/value – Collecting user-generated content is a novel research issue in the library sciences field, and few studies have developed useful tools that allow readers to actively contribute their knowledge to digital libraries. This work shows how to implement such digital library systems and how the annotation tool benefits the growth of digital archives and promotes learning performance.

Keywords Digital libraries, Annotation, DSpace, Knowledge archiving, Knowledge sharing, Knowledge management, Taiwan

Paper type Technical paper
1. Introduction
Libraries are important institutions that preserve human culture. Beginning in the 1990s, the Internet has contributed markedly to the development of digital libraries (DLs). Some predict that DLs will be the predominant libraries in the future. Library and education professionals have recently become actively involved in DL research and development. Many DL projects (e.g., American Memory of the Library of Congress (http://memory.loc.gov/ammem/index.html) and the Digital Archives Project of National Taiwan University (www.darc.ntu.edu.tw/)) focused on preserving human cultural heritage, historical relics, and institutional productions. However, few DL projects have focused on preserving library history. Taiwan’s library history can be divided into the following four periods: the Japanese occupation period (1895-1945); the reconstruction period (1945-1952); the growth period (1953-1976), and the rapid growth period (1977-present). Many historical events are valuable and should be collected in a DL to support teaching and academic research in the library and information field. This work developed a DL for Taiwan libraries’ history to preserve the developmental history of Taiwan’s libraries by modifying the open-source DSpace platform.

2. Background
Digital libraries are focused on creating, organizing, and managing multimedia digital content and collections, and providing search, retrieval, and other information services over computer networks. Currently, some open-source software packages (Goh et al., 2006; Jody, 2007), such as Eprints, Fedora, Greenstone, GERN CDSware, and DSpace, are available to organizations and individuals to create DLs. Open-source DLs are increasingly considered a beneficial alternative to commercial DL systems due to user dissatisfaction with the functionalities of commercial DL systems and the increasing budget cuts libraries now face. Additionally, the costs of maintenance and producing software are also increasing dramatically. Free access and a good level of functionality are the main reasons accounting for the usage and interest in open-source DL software.

Goh et al. (2006) developed a checklist consisting of 12 criteria for evaluating open-source DL software. The 12 features that characterize “good” open-source DL software are effective content management and content acquisition, the existence of metadata, search support, adequate access control and privacy, report and inquiry capabilities, preservation of digital content, interoperability, a user-friendly interface, standards compliance, effective automated tools, and good support and maintenance. Although Goh et al. (2006) evaluated four open-source DL software programs – Eprints, Fedora, Greenstone, and GERN CDSware – their study did not analyze DSpace (www.DSpace.org). The first version of DSpace was released in November 2002, following a joint effort by developers from Massachusetts Institute of Technology (MIT) and Hewlett-Packard (HP) Labs in Cambridge, Massachusetts. Notably, DSpace has powerful functionalities that are characteristics of a good DL, as identified by Goh et al. (2006) in their study of archiving, retrieval, and management of digital assets. DSpace is currently used as a basis for institutional repositories (Tansley et al., 2003).

Although DSpace has sufficient functionalities for archiving, searching, and managing digital content, the predefined metadata schema needs to be revised to archive digital targets, such as library history in Taiwan. Additionally, the user
interface in DSpace for browsing digital content is very primitive and clearly cannot satisfy user needs. More importantly, the original DSpace system lacks an interface for exhibiting various digital objects, such as images, video, audio, and animation. Moreover, collecting content for DLs is the most important and difficult step in developing rich digital content. Recently, Web 2.0 technologies have enabled the contribution of user-generated content to DLs and developing new applications with Web 2.0 technologies in a library setting is becoming an important research issue.

Curran et al. (2007) noted that user-generated content, as opposed to content posted solely by site author(s), and treating users as co-developers of a site are features of Web 2.0 websites. Wikipedia is the most well-known site for user-generated content. Wikipedia, the largest multilingual free encyclopedia on the Internet, allows any user to add content, and all content can be edited by any other user. The principles of Web 2.0 that are applicable to DLs are interactivity and the possibility for users to contribute content. Many studies have indicated that users can create additional information to digitalized objects in DLs. In developing Web 2.0 technologies to DLs, the Oregon State University libraries integrated Reference Desk Manager (RDM) with a Wiki to enhance their reference services (Frumkin, 2005). Based on a literature survey, Web 2.0 services that help users create strong relationships between users and a library are urgently needed.

At present, Wiki, blogs, social tagging, Really Simple Syndication (RSS), social networks, and annotations are the main Web 2.0 technologies that have high potential for application to DLs. In particular, few studies have applied annotation technology to print libraries or DLs. Readers typically annotate printed books as a routine part of their engagement with materials. Marshall (1997) proved that annotations in books are useful to subsequent readers. Traditionally, annotating printed books by pen is the most common method of recording book knowledge, but is disadvantageous to knowledge storage in computers, knowledge dissemination, and knowledge sharing via the Internet. In contrast, annotating digital documents via an annotation tool can overcome these shortcomings. Petri et al. (2005) categorized annotation systems as document-centered and discussion-centered systems. Discussion-centered systems provide tools for browsing and annotating discussions, while document-centered systems allow learners to browse and annotate web-based documents. Frumkin (2005) indicated that if users could leave comments or annotations on a finding aid – providing additional information related to materials – this practice would open the door to sharing research experience, facilitate collaborative research, and make it easy for future researchers to find materials they need in a particular collection. Additionally, many studies (Ovsiamnikov et al., 1999; Petri et al., 2005; Rau et al., 2004; Hwang et al., 2007) demonstrated that applying annotation tools to aid learning promotes learning performance. Constructing rich digital content and users leaving knowledge in DLs via annotations are important research issues.

This work presents a novel Web 2.0-based reader knowledge archiving and sharing tool. By using this tool, readers can conveniently add annotations to archived digital materials in Taiwan library history DL (http://tlh.lias.nccu.edu.tw). The functionalities of the proposed annotation tool allow readers to add explanations, synonyms, comments, and hyperlinks for marked target areas, and underline paragraphs as important sections as well as rate annotations from other users. With the proposed
annotation tool, annotations for archived digital material can be accumulated continuously by readers annotating the material. A digital material archive with rich annotation information can increase the reading comprehension of first-time readers as well as promote deep reading for readers who have read the material. Moreover, annotation information also promotes reading efficiency because some additional comments attached to the digital material provide meaningful and interpretive information that can shorten reading comprehension time. In conclusion, this work explores the effects of cooperative annotation – its value to annotators and subsequent readers – and the impact of this practice has on the Taiwan library history DL.

The remainder of this paper is organized as follows. The development of a DL that preserves Taiwan library history using the open-source DSpace platform is presented in section 3. The proposed Taiwan library history DL with the annotation tool for reader knowledge archiving and sharing is described in section 4. Section 5 assesses the effect of this proposed tool on reader performance. Finally, conclusions and directions for future work are given in section 6.

3. Developing a DL for Taiwan’s library history using the DSpace platform

Based on these shortcomings of DSpace system mentioned in section 2, this work modified and enhanced the functionalities of the original DSpace system to develop a Taiwan library history DL. The modifications of functionalities are described as follows:

1. **Modifying the predefined metadata schema.** DSpace predefines Dublin Core as metadata schema for describing a digital object. Many institutional repositories employ Dublin Core directly as a metadata schema to archive targeted digital content. However, it is insufficient for developing a Taiwan library history DL because this DL must archive complex digital objects, including library personnel, library architecture, library information, educational content, library timelines, and library archives. Figure 1 shows a diagram of the designed metadata structure for archiving Taiwan library history. Basically, DSpace can define custom metadata schema for a new item by modifying the metadata form definition. The detailed modification processes are explained in the technical document entitled Custom Metadata-entry Pages for Submission (http://scm.DSpace.org/svn/repo/sandbox/DSpace-dao-prototype/DSpace/docs/submission.html#metadataEntry). Figure 2 shows a modified XML file and the corresponding metadata submission entry for library personnel. Currently, six metadata schemas are designed for the Taiwan Library history DL.

2. **Modifying the language interface.** DSpace predefines English as the main language for DLs, and also provides a language modification interface that supports non-English users. The detailed modification processes are explained in DSpace 1.4.1 Beta 1 System Documentation (https://wiki.duraspace.org/download/attachments/19006155/DSpaceStandard141beta1.pdf?version=1&modificationDate=1283190822509). Figure 3 shows the developed language transformation program for transforming Chinese words with the Big 5 code into Unicode. For example, the Chinese term 「台灣百年圖書館史」is transformed into [u53f0]u7063[u767e]u5e74[u5716]u661f[u9928]u53f2] by this program.
Modifying the user interface. The original DSpace user interface is not easy to use. Fortunately, DSpace also provides a modification mechanism that permits improvements to the user interface. The detailed modification processes are also explained in DSpace 1.4.1 Beta 1 System Documentation. To modify the DSpace user interface, with the exception of database operation, one must be familiar with HTML, Cascading Style Sheets (CSS), JavaScript, JavaServer Pages (JSP), and JAVA programming languages. Figure 4 shows the modified DSpace user interface for the Taiwan library history DL. The modified DSpace user interface

Figure 1.
The diagram of metadata structure for archiving Taiwan libraries' history
Modifying XML file for setting a corresponding metadata submission entry for library personnel
Figure 3. Language transformation program for transferring Chinese to Unicode

Figure 4. The modified DSpace user interface
has six parts – a header, a location bar, a navigation bar, content, a sidebar, and a footer.

Next, the function of each display area in the DSpace user interface is explained as follows:

- **Header area.** The header area displays the homepage logo. In addition to displaying the homepage logo, this area also shows the final date that uploads digital object, and the total number of digital items archived in the Taiwan library history DL.

- **Location bar.** The location bar can display the complete hierarchical layer names for users browsing from the homepage to their current browsing location. The bar also provides corresponding hyperlinks, such that users can conveniently jump to any hierarchical layer they have surfed. Another important function stops users from losing information in the hyperspace of the Taiwan library history DL.

- **Navigation bar.** The navigation bar allows users to click target items rapidly while browsing. The bar comprises user and administrator interfaces.

- **Content area.** The content area is the main area displaying digital items on which a user clicks. This area is controlled by JSP markup language.

- **Side bar.** Compared with the navigation bar, the side bar can be considered the second navigation bar, which can display some assistive information such as a website map, RSS services, and website instructions.

- **Footer area.** The footer area displays such information as contact and copyright information.

(4) **Designing an interface for exhibiting various digital objects.** In the original DSpace system, if a user wants to browse an archived digital object, then the file of that digital object must be downloaded to the user’s computer from the DSpace system. In other words, DSpace does not have an interface for displaying various digital objects, such that opening a digital object for viewing must utilize the software installed on a user’s computer. To develop an excellent digital library, an exhibition interface is needed to present various digital objects. Particularly, the Taiwan library history DL needs a powerful presentation interface with such functionalities as enlarging, shrinking, and dragging an image because the library contains many rare photos, important archives, and old newspapers associated with library architecture, library personnel, library education, and library organization. To solve this problem, this work integrated image browser Zoomifyer EZ (www.zoomify.com/express.htm), which is an open-source software program that uses Flash technology for exhibiting images on Internet browsers with the DSpace system. Figures 5 and 6 show the user interface developed using Zoomifyer EZ software exhibiting the library building at National Taiwan University and a newspaper archived in the library at National Chengchi University, respectively.
**Figure 5.**
Zoomifyer EZ interface for exhibiting the building of National Taiwan University Library.

**Figure 6.**
Zoomifyer EZ interface for exhibiting old newspaper.
4. Taiwan library history digital library with the annotation tool for archiving and sharing reader knowledge

4.1 System architecture

Figure 7 presents the system architecture of the proposed annotation tool in the Taiwan library history DL for archiving and sharing reader knowledge. The proposed DL with the user annotation tool consists of the Taiwan library history DL, the annotation module, an annotation interface, and two databases. The DL provides materials readers browse, search, read, and annotate. The annotation service module integrates original material archived in the DL with annotations from readers to form new DL material with additional information and/or contexts. To avoid destroying the original version of archived materials, the annotation service module uses an independent annotation database, which differs from the DL database, to store annotated information and contexts from readers. Moreover, the annotation interface provides several powerful annotation functionalities, such as annotating key paragraphs in a text via underlining, explaining unknown terms, providing comments on marked words or sentences, and assigning hyperlinks to marked words or sentences, to assist readers annotating an original archived material in the Taiwan library history DL. The annotation interface also has a knowledge-sharing interface that facilitates knowledge exchange among readers who read annotated materials in the DL.

4.2 System operation procedure

Based on the system architecture, the system operating procedure is described and summarized as follows.

A user requests a webpage archived in the Taiwan library history DL for reading by assigning a URL for browser. The DL system then accesses the webpage that the user requested from the DL database.

The DL system shows the original webpage in HTML for reader annotation. The DL system then passes the URL address given by the user to the annotation service.

![Figure 7. System architecture of the proposed annotation tool on Taiwan library history digital library](image-url)
The annotation service module accesses annotation data on the webpage from the annotation database. The annotation service module integrates the original webpage and annotation data for this webpage and stores it in an annotated webpage using Java script technology. The integrated webpage with the original webpage and annotation data is passed to the annotation interface, such that a user can read the annotations.

4.3 The functionalities of the user annotation tool in the Taiwan library history digital library for archiving and sharing reader knowledge

Figure 8 shows the proposed reading annotation tool in the Taiwan library history DL. To avoid interference with usage of the DL, the annotation tool can be hidden as a small icon located in the upper-right portion of the screen when no annotation is needed. If a reader wants to know which words are annotated on a digital material, they can select the “browser” menu, which highlights all words in the annotation. Figure 9 displays the highlighted target areas with annotations. Additionally, users must first register to become legal users via a registration wizard before using the reading annotation tool. Once a user logs on to the annotation tool, he/she will be able to use all annotation functionalities, including highlighting and adding comments. That is, users can annotate digital material in the Taiwan library history DL by highlighting and commenting on selected target areas. Figure 10 shows the menu for selecting an annotation type to annotate a selected target area. To annotate a selected target area, a user can select various annotation types, including explanation, synonym, comment, hyperlink, and underline. Users can browse all annotations for digital material on the DL. Figure 11 shows this user interface. A target area with multiple annotations is open, allowing users to communicate asynchronously for sharing knowledge via annotations. Therefore, each annotation can be shared by all system users, and clicking on an annotation provides detailed information such as annotation author, rating result, and comments.

Figure 8.
Reading annotation tool on Taiwan library history digital library
Moreover, users can also edit their own annotations; that is, annotations generated by other users cannot be edited, they can only be given a rating. Rating an annotation is helpful in identifying high-quality annotations. The proposed annotation tool can display annotated content with the highest ratings when a mouse stays on a selected target area with an annotation (Figure 12). Of course, users can also click on a selected target area with annotations when browsing annotations from all readers; Figure 13 shows this user interface. Finally, users can also provide summary comments and rate...
the quality level for annotated digital materials on the Taiwan library history DL; Figure 14 shows this user interface.

5. Assessing the enhancement of reading performance
The work conducted an experiment to assess whether the Taiwan library history DL with the proposed annotation tool is superior to the Taiwan library history DL without the proposed annotation tool in terms of user reading performance. In this experiment, 70 undergraduates studying at National Chengchi University were invited to participate. Among the 70 undergraduates, 35 served as the control group and used the Taiwan
library history DL without the proposed annotation tool; the remaining 35 students served as the treatment group and used the DL with the proposed annotation tool. Both groups completed a pretest and post-test for comparing differences in reading performance before and after learning. In this experiment, a subject related to “Taiwan library personnel” on the Taiwan library history DL was selected as the reading target. Before learning, all participators were invited to write a report based on prior knowledge of the selected subject. Next, the two groups began the reading activity to understand the history of Taiwan library personnel based on the digital content archived in the Taiwan library history DL. After learning, all participants were invited to revise their reports finished during the pre-test as their reading outcomes. The finished reports were evaluated by the course instructor. Table I shows descriptive statistics of learning performance for both the control and treatment groups.
First, the independent samples t-test was applied to assess whether prior knowledge of two participating groups for “Taiwan library personnel archived in the Taiwan library history DL was equivalent before the experiment based on pre-test results. Analytical results indicate that pre-test scores for these two groups did not differ significantly (Sig. = 0.2006 > 0.05). Therefore, prior knowledge of two participating groups can be viewed as equivalent before the experiment. Next, post-tests of the two participating groups were also analyzed using the independent samples t-test. The t-test result (Sig. = 0.0424 < 0.05) shows that the post-test scores of the treatment group were significantly higher than those of the control group. Moreover, this study further compared pre-test and post-test scores for each group using the paired samples t-test. Analytical results for the control group indicate that the difference in mean pre-test and post-test scores is −1.56, and the paired samples t-test result reaches the significance level (Sig. = 0.022 < 0.05). Additionally, the pre-test and post-test scores for the treatment group were also assessed using the paired samples t-test. Analytical results indicate that the difference mean pre-test and post-test scores is −2.96, and the paired samples t-test result (Sig. = 0.000 < 0.05) shows that the treatment group also achieves significant difference. Experimental results show that both groups had improved learning performance, but the treatment group improved more than the control group.

6. Conclusions and future work

Popular research on DLs mainly focused on organizing and accessing information, and most ignored the importance of archiving reader knowledge. Since reader knowledge cannot be immediately stored and efficiently managed, DLs lose a large amount of useful reading outcomes and processes from readers who use DLs. Thus, reader knowledge cannot be applied or efficiently shared through DLs or the Internet. Therefore, this work presented a novel tool for archiving and sharing reader knowledge that allows readers to easily annotate digital materials in HTML in the Taiwan library history DL, thereby enabling cooperative knowledge creation. Experimental results show that the proposed archiving and sharing tool effectively gathers reader knowledge via user participation. Compared to the original digital material without annotations, digital material with annotations, such as explanations, user comments, and rich contexts associated with digital materials, is very useful to

<table>
<thead>
<tr>
<th>Comparison item</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of learners</td>
<td>Mean</td>
<td>Std. deviation</td>
</tr>
<tr>
<td>The control group performed the reading learning without the support of the reading annotation tool</td>
<td>35</td>
<td>75.57</td>
</tr>
<tr>
<td>The treatment group performed the reading learning with the support of the reading annotation tool</td>
<td>35</td>
<td>76.41</td>
</tr>
</tbody>
</table>

Table I. The descriptive statistics information of reading learning performance for both the control and treatment groups.
developers. The values to those annotating and subsequent readers are the acquisition of deep knowledge and increased reading efficiency. Furthermore, the implications of the work are that DL content grows dynamically as readers contribute knowledge, thus assisting in the development of a reading community on DLs. A more important issue is that annotated information from different readers has very high potential for value-added reading knowledge utilizing data mining techniques.

Although this work has presented several implications of the proposed tool for archiving and sharing reader knowledge in DLs, several other issues warrant further investigation. Categorizing annotations as public and private will be a significant issue in the future. Public annotations are for sharing knowledge, and private annotations can preserve personal reading outcomes for knowledge management. Providing a discussion board for comments or replies to annotations from readers can increase interaction among readers. This would also prove helpful in creating a learning community for a DL. Ming reading knowledge based on reader annotations with ratings is a potentially fruitful research direction. Finally, the Taiwan library history DL with the proposed annotation tool may benefit e-learning.

References


About the authors

Chih-Ming Chen is currently an Associate Professor in the Graduate Institute of Library, Information and Archival Studies at National Chengchi University, Taipei, Taiwan. He received BSc and MSc degrees from the Department of Industrial Education at National Taiwan Normal
University in 1992 and 1997, and received his PhD degree from the Department of Electronic Engineering at National Taiwan University of Science and Technology in 2002. His research interests include digital libraries, e-learning, data mining, machine learning and intelligent agents on the web. Chih-Ming Chen is the corresponding author and can be contacted at: chencm@nccu.edu.tw

Yong-Ting Chen received her BSc. degree from the Department of Information and Library Science at Fu Jen Catholic University in 2007, and is now a Graduate Student of the Graduate Institute of Library, Information and Archival Studies at National Chengchi University. Her research interests include digital library and e-learning.

Chin-Ming Hong is currently a Professor in the Department of Applied Electronics Technology at National Taiwan Normal University, Taipei, Taiwan. His research interests include intelligent control, fuzzy theory, and home health care.

Chin-Wen Liao is currently an Associate Professor in the Department of Industrial Education and Technology at National Changhua University of Education, Changhua, Taiwan. His research interests include power education, innovation management, industrial education, industrial design and automatic control.

Chia-Meng Huang is currently a software engineer in ADLINK Technology Inc., Taiwan. His research interests include embedded system and physiology signal processing.